Owen, G., Jones, K., & Harris, R. (2017). Does neighbourhood deprivation affect the genetic influence on body mass? *Social Science and Medicine*, *185*, 38–45. https://doi.org/10.1016/j.socscimed.2017.05.041

Does Neighbourhood Deprivation affect the Genetic Influence on Body Mass?

Abstract

Most research into the role of gene-environment interactions in the etiology of obesity has taken environment to mean behaviours such as exercise and diet. While interesting this is somewhat at odds with research into the social determinants of obesity in which the focus has shifted away from individuals and behaviours to the types of wider obesogenic environments in which individuals live, which influence and produce these behaviours. This study combines these two strands of research by investigating how the genetic influence on body mass index (BMI), used as a proxy for obesity, changes across different neighbourhood environments measured by levels of deprivation. Genetics are incorporated using a classical twin design with data from Twins UK, a longitudinal study of UK twins running since 1992. A multilevel modelling approach is taken to decompose variation between individuals into genetic, shared environmental and non-shared environmental components. Neighbourhood deprivation is found to be a statistically significant predictor of BMI after conditioning on individual characteristics; and a heritability of 0.75 is estimated for the entire sample. This heritability estimate is shown, however, to be higher in more deprived neighbourhoods and lower in less deprived ones and this relationship is statistically significant. While this research cannot say anything directly about the mechanisms behind the relationship it does highlight how the relative importance of genetic factors can vary across different social environments and therefore the value of considering both genetic and social determinants of health simultaneously.

Keywords

Gene-environment interactions

Neighbourhood deprivation

Multilevel modelling

Obesity

Introduction

Obesity is an important public health issue due to its links to chronic diseases such as type 2 diabetes and hypertension (Kopelman, 2007) and association with increased mortality generally (Flegal et al., 2013). The worldwide prevalence of obesity has more than doubled since 1980 with the World Health Organisation (2015) describing the problem as an "escalating global epidemic". In recent years lack of success with interventions at the individual level in reversing this trend has led to a substantial amount of research aimed at understanding the wider food and built environments in which people live and work which promote obesity's proximate causes of poor diet and sedentary behaviour (Lake and Townshend, 2006). These environments are often labelled 'obesogenic' (Egger and Swinburn, 1997). While there has been a growing body of literature on how these types of environments affect obesity, one area that has so far been unexplored is how these environments may affect the genetic influence on obesity.

Obesity has been shown to have a genetic influence. Studies exploiting the genetic relatedness of twins have estimated the heritability of body mass index (BMI), defined as the proportion of variation in a trait attributable to variation in genetics, as anywhere between 0.47 and 0.9 depending on the population studied and the method used (Elks et al., 2012). It has become increasingly acknowledged, however, that the heritability of many human traits such as BMI, is not a constant and is dependent on the social environment (Turkheimer et al., 2003, Tuvblad et al., 2006). In other words, social factors and genetics may interact in producing health outcomes and as such cannot be assumed to be independent. Despite this there has been little research into how wider social contexts such as neighbourhoods may moderate genetic influence on obesity. Previous gene-environment interaction studies have mostly treated the environment as behaviours such as exercise (e.g. Ahmad et al. (2013)) and diet (e.g. Qi et al. (2014)) rather than the wider obesogenic environments which produce these behaviours and which are currently the focus of much social research (Boardman et al., 2013). Notable exceptions include a study by Boardman et al. (2012)

which shows that school context moderates heritability of body mass index (BMI) and Rosenquist et al. (2015) who use measured genetic data to show that the relationship between variants of the fat mass and obesity associated FTO gene and body mass index varies with cohort of birth.

Understanding the relative importance of genetic and social risk factors and how they interact can be important for future health policy. For example it may be the case that in the most extreme obesogenic environments, the environment 'outweighs' the effect of an individual's genetics, having such a large effect that differences between people due to genetics are small in comparison. This may not be a causal interaction in the biological sense as the change in heritability would be due to an increase in environmental variation rather than changes in the effect of specific genes. If this were the case it would suggest that more resources should be put into focussing on understanding social environments.

Alternatively it may be that there are aspects of obesogenic environments that moderate genetic vulnerability to obesity and that cause changes in the relationship between specific genes and obesity. If this was the case, policies could be designed that not only reduce the effect of the environment as a whole but can also reduce the genetic influences on weight gain (Boardman et al., 2012). Another possibility could be that social policies are only effective for certain types of individuals due to their genotype. Whatever the underlying relationship between social environments and genetics, these possibilities highlight the need for research which considers both genetics and the social environment together.

In this research, neighbourhood deprivation is used as the proxy for the obesogenic environment as, according to a systematic review from Giskes et al. (2011), it is the only measure of neighbourhood environment consistently associated with obesogenic dietary intakes, even after controlling for possible confounding individual level variables. The theory is that deprived neighbourhoods may be poorer environments for food and physical activity, though the empirical evidence for the pathways through which this happens is mixed, and may differ across countries (Cummins and Macintyre,

2006, Townshend and Lake, 2009). One possibility is that access to unhealthy food is easier in deprived neighbourhoods. For example Cummins et al. (2005) show that in the UK, neighbourhood deprivation is correlated with the number of McDonalds restaurants. Another possibility is that more deprived neighbourhoods have poorer quality recreational facilities and greenspaces which may discourage physical activity (van Lenthe et al., 2005), although other research suggests access to facilities for physical activity can actually be better in deprived neighbourhoods (Pearce et al., 2007).

Additionally, as well as the built and physical environments it may be that the social environments of neighbourhoods matter too. Neighbourhoods may have an influence on health behaviours through their influence in forming social norms and through social networks (Galster, 2012). Living in deprived neighbourhoods may also contribute to increased stress, which has been suggested to have an influence on increasing unhealthy behaviours (Pampel et al., 2010). Finally, although neighbourhoods are the focus of this study, it is important to remember that neighbourhoods themselves are only one aspect of obesogenic environments. Larger scale political and economic contexts and other small scale contexts such schools may also play an important role (Swinburn et al., 1999).

The aim of this research is to examine whether the genetic influence on BMI, varies as a function of neighbourhood deprivation and if so, what the nature of this relationship is. The research is based upon data from the UK and genetic influence is included latently as heritability using data from twin pairs. The paper will proceed by first introducing the data, then outlining the statistical methodology before presenting the results and discussing their implications.

Data

The study uses data from Twins UK, an ongoing study of twins aged 16 and over from across the UK which began in 1992. Twins chose to become part of the study, meaning that the sample is not representative of the UK population. For historical reasons the sample is 90% female. Additionally, the sample is more highly educated and has proportionally fewer ethnic minorities than the UK

population as whole. The dataset consists of measurements of height and weight, from which BMI can be calculated. These measurements were taken at various points between 1992 and 2007, as different individuals joined the study at different times and there were different waves of measurement. Some individuals were measured more than once and if this was the case the most recent measurement was used for the analysis shown in the results.

Geographical identifiers for the twins were available as postcode sectors but, as there are few statistics calculated at this spatial scale, deprivation data was taken from 2001 Carstairs scores calculated at the ward level. Wards contain on average around 6,600 individuals while postcode sectors contain approximately 5,000 and often coincide with parts of multiple wards. Wards therefore had to be matched to postcode sectors. This was carried out using the UK data service *geoconvert* tool (http://geoconvert.mimas.ac.uk/) which additionally gave the proportion of each postcode sector by area that belonged to each ward. From this a deprivation score was calculated for each postcode sector using a weighted average by area of all the wards that overlapped with that postcode sector. For 19% of the twin pairs, the two twins lived in the same postcode sector, although this was more common among younger twins. The correlation in terms of neighbourhood deprivation between co-twins was 0.45.

BMI is included as the dependent variable, however the natural logarithm is used as residual diagnostics of preliminary analyses showed that the residuals were not normally distributed. Age, sex, ethnic origin, and education as a proxy for socioeconomic status were used to control for individual level confounding as these variables may have an effect on both selection into neighbourhoods and BMI (Diez Roux and Mair, 2010). Additionally a control was added for year of measurement, as measurements were taken at different points in time and, during the time of the study, there was an increase in obesity prevalence at the population level.

The original dataset contained data on 7,629 individuals. Twin pairs in which one or both did not have a single valid BMI reading were excluded, along with pairs with uncertain zygosity and twins who had been raised separately. Zygosity refers to whether the twin pair are monozygotic (identical) or dizygotic (fraternal). Twin pairs who had missing data on residential location or on one of the control variables were also dropped from the analysis. This left 3,128 individual observations consisting of 830 pairs of monozygotic twins and 734 pairs of dizygotic twins. As far as it was possible to assess there appeared to be no large systematic differences between the cases with complete data and the cases that had missing data. Summary statistics and information about the variables are presented in Table 1.

Methodology

This study uses the classical twin design which is based around comparing the correlations between monozygotic (MZ) and dizygotic (DZ) twins. The correlation between MZ twins and their co-twins is made up of environmental influences, that the twins share, and genetic influences. The correlation between DZ twins and their co-twins is similarly made up of shared environmental influences and genetic influences but in contrast to MZ twins who share 100% of their genes, DZ twins share on average only 50% of their polymorphic alleles (genes that occur in more than one form). This means, assuming that shared environmental influences on MZ twins and DZ twins are the same, the differences between the correlations between MZ and DZ twins will be due to 50% of the genetic influences. From this information it is also then possible to partition the remaining variation into shared environmental and non-shared environmental components.

Put algebraically:

$$r_{mz} = A + C$$

$$r_{dz} = \frac{1}{2}A + C$$

$$A = 2(r_{mz} - r_{dz}) \quad (1)$$

$$C = r_{mz} - A$$

$$E = 1 - r_{mz}$$

Where r_{mz} is the correlation between MZ twins, r_{dz} is the correlation between DZ twins, A is the heritability estimate or the proportion of variation due to genetic influences and C and E are the proportions of variation attributable to shared and non-shared environmental influences respectively. The shared environment is defined here as anything which makes the twins more similar which is not attributable to genetics, and is likely to include the effects of family upbringing and socioeconomic characteristics. This could also include the effect of the neighbourhood environment if the twins live in the same or similar neighbourhoods. The non-shared environment incorporates any environmental influences which make twins different from their co-twin and which may also include measurement error.

It should be noted that this method makes a number of assumptions which have been debated widely and still remain contentious (Beckwith and Morris, 2008). Firstly it assumes that variation attributable to the shared environment is on average the same for MZ twins as for DZ twins. This is known as the equal environments assumption and may be broken if, for example, people treat MZ twins on average more similarly due to their more similar appearance. Other assumptions are random mating, meaning that people are as likely to choose partners who are genetically similar to those who are genetically different, and that the genetic mechanisms involved are additive. Finally

there is the issue of whether findings from twins can be generalised to non-twins, who make up the majority of the population.

While it is possible to calculate an estimate of heritability simply by using the correlations, use of a modelling strategy allows for testing of varying hypotheses, and for the heritability of BMI to be conditional on neighbourhood deprivation. Using procedures outlined by Rabe-Hesketh et al. (2008) standard multilevel modelling techniques can be adapted to partition the variation in BMI into genetic, shared environmental and non-shared environmental components.

A multilevel heritability model can be written:

$$y_{ij} = \beta_0 + a_{0ij} + c_{0j} + e_{0ij}$$

$$a_{0ij} \sim N(0, \sigma^2_{a0})$$

$$c_{0j} \sim N(0, \sigma^2_{c0}) \qquad (2)$$

$$e_{0ij} \sim N(0, \sigma^2_{e0})$$

$$\text{cov}(a_{01j}, a_{02j}) = r\,\sigma^2_{a0}$$

Where y_{ij} is the response, in this case log BMI, for individual *i* in twin pair *j*, β_0 is the grand mean log BMI and $a_{0ij}, c_{0j}, \text{and } e_{0ij}$ are random effects whose variances σ^2_{a0}, σ^2_{c0}, and σ^2_{e0} are the amount of variation in log BMI attributable to additive genetic, shared environmental and non-shared environmental variation respectively. *r* is the relatedness for twin pair *j* (1 for MZ twins or 0.5 for DZ twins). The model has two levels, with individuals at level 1 and twins at level 2, but three random effects, for the three distinct types of variation which are specified and estimated. A heritability estimate can be calculated by dividing the variance due to genetics, σ^2_{a0}, by the total variance.

$$heritability = \frac{\sigma_{a0}^2}{\sigma_{a0}^2 + \sigma_{c0}^2 + \sigma_{e0}^2} \quad (3)$$

This model can be extended to allow each variance component to itself vary, as a function of a measured environmental variable, referred to as a differential heritability model (Pillinger, 2012). This then enables the inclusion of a gene-environment interaction as in the model genes are allowed to 'interact' with the environment, with heritability dependent on social context. Following Pillinger (2012) this model can be specified:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + a_{0ij} + c_{0j} + e_{0ij} + a_{1ij} x_{1ij} + c_{1j} x_{1ij} + e_{1ij} x_{1ij}$$

$$\begin{bmatrix} a_{0ij} \\ a_{1ij} \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma_{a0}^2 & \\ \sigma_{a01} & \sigma_{a1}^2 \end{bmatrix}\right)$$

$$\begin{bmatrix} c_{0j} \\ c_{1j} \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma_{c0}^2 & \\ \sigma_{c01} & \sigma_{c1}^2 \end{bmatrix}\right) \quad (4)$$

$$\begin{bmatrix} e_{0ij} \\ c_{1ij} \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma_{e0}^2 & \\ \sigma_{e01} & \sigma_{e1}^2 \end{bmatrix}\right)$$

This is the same as equation 2 except that x_1, a measured environmental variable (neighbourhood deprivation in the case of this research) has now been added to the model and a second random effect associated with each component has been added that is dependent on x_1. The two random effects for each component are assumed to come from a multivariate normal distribution, the variance of which is a quadratic function of x_1. For example, the variance attributed to genetics can be expressed as $\sigma_{a0}^2 + 2\sigma_{a01} + \sigma_{a1}^2$. It is however, possible to specify other forms of functions by constraining a combination of the variance or covariance parameters to 0. Likelihood ratio tests are used to determine the functional forms of the variances that best fit the data.

To answer the research question 'does neighbourhood deprivation affect genetic influence on body mass', a set of models are built and estimated using the iterative generalised least squares algorithm in MLwiN 2.33. Model 1 is a multilevel model for calculating heritability as in equation 1, adjusted for age, year and sex. This allows a heritability estimate to be calculated which can be compared with previous studies. Model 2 is the same as Model 1 but with the inclusion of neighbourhood deprivation as a predictor variable alongside the individual and twin level control variables, to assess whether neighbourhood deprivation is associated with log BMI. Model 3 is the differential heritability model, again including these control variables, but allowing heritability to vary as a function of neighbourhood deprivation.

Results

Table 2 shows the results of all models. Using equation 3 to calculate the proportion of the total variation due to genetics, Model 1 estimates a heritability of 75% across the sample. The majority of the rest of the variation is attributed to the non-shared environment, while the variance attributed to shared environmental factors is small and not statistically significant. Model 2 shows that there is a statistically significant association between neighbourhood deprivation and log BMI, even after controlling for the possible confounding individual variables of education and ethnic origin.

Model 3 shows that the heritability of BMI varies as a function of neighbourhood deprivation. As the variance and covariance parameters have little meaning in isolation, the results have been displayed graphically in Figures 1 and 2 for ease of interpretation. The heritability estimate can be calculated from the proportion of variation that is due to genetics across all levels of deprivation. Figure 1 displays this estimate and shows that heritability is generally smaller in less deprived neighbourhoods than in more deprived neighbourhoods. This correlation between increased heritability and increased deprivation is statistically significant although the relationship is non-linear and the trend levels off, before heritability starts to decline in the most deprived environments. Because fewer than 10% of individuals live in a neighbourhood with a deprivation score greater than

4, the confidence intervals are wide for the most deprived neighbourhoods, indicating substantial uncertainty around this decline. At least, however, it seems possible to say that for neighbourhoods in the upper quartile of deprivation there is no additional increase in heritability for additional deprivation.

Figure 2 shows how the actual amounts of variation change for all three of the variation types. Although the absolute scores on the y axis are less meaningful, this is useful for two reasons. Firstly it shows absolute changes in variation due to genetics rather than changes in the proportion of variation due to genetics relative to environmental variation. The change in heritability shown in Figure 1 could be solely due to changes in the importance of the environment. Figure 2, however, shows that this is clearly not the case and that there is an increase in the amount of variation due to genetics for people in more deprived neighbourhoods. Secondly it shows how environmental influences vary across neighbourhoods. Non-shared environmental influences appear to be relatively similar across the spectrum of neighbourhoods, however shared environmental influences appear only to be important in less deprived neighbourhoods and disappear in more deprived neighbourhoods.

Discussion

The results of this study show that the genetic influence on obesity is dependent upon the social environment with a higher proportion of the variation in BMI being attributed to genetics in more deprived neighbourhoods. Additionally the results suggest that this increase in heritability is at least in part due to the actual amount of variation linked to genetics increasing, rather than solely due to changes in the amount of environmental variation. The results show no evidence for environmental influences taking over and 'outweighing' the genetic influences in the most obesogenic environments. Instead, the genetic influences actually become stronger. This indicates some kind of social trigger or social control mechanism in which either something associated with living in a deprived neighbourhood triggers genetic vulnerability for obesity or, alternatively, something

associated with less deprived neighbourhoods supresses it (Boardman et al., 2012). This concurs with the idea of Loos and Bouchard (2003) who suggest that some individuals may have a slight genetic predisposition to gain weight but that this only manifests itself when individuals are subject to obesogenic environments which encourage poor diet and lack of physical activity.

The key message of this study is that genetic associations are not constant over place and social context and, therefore equally the effects of the social environment will not be the same for all individuals due to differences in their genetics. While this is certainly not a new idea (Turkheimer et al., 2003, Shanahan and Hofer, 2005) this has profound policy implications which have, thus far, been relatively unexplored, as the effect of social policies, including those at the level of the neighbourhood, are likely to be moderated by an individual's genotype. In this case as genetic influence appears to be higher in deprived environments, the goal would be to search for social policies which can reduce genetic vulnerability as well as social vulnerability in these environments.

It is important, however not to overstate the findings of this study. Much more research is required to truly understand the mechanisms behind the gene-environment interaction before this knowledge can be fed into effective policy making. Firstly, in this study, genetics have been included latently as heritability so this research cannot provide information on the possible genetic mechanisms. While heritability is still a useful concept, as it provides a good general view of genetic risk (Dick, 2011), this type of study needs to be complimented by genome wide studies based on measured genetics.

Secondly, it is not possible to claim a causal link between neighbourhood deprivation and changing genetics. This is because though the analysis conditioned on possible confounding individual variables in the fixed part of the model, the model did not control for possible confounding gene-environment interactions (Keller, 2014). For example the change in heritability may be due to some individual social characteristic which is associated with living in a deprived neighbourhood. Further research with large sample sizes, therefore, needs to be undertaken to test for many different

possible gene-environment interactions and in this the choice of environmental variables must utilise prior knowledge gained from research into obesogenic environments.

This study also has a number of other limitations. Due to data availability the neighbourhood used in the analysis is defined by the postcode sector. This is potentially problematic as postcode sectors are quite large (with an average of around 5,000 inhabitants) and were created for logistical reasons, not for the purpose of representing neighbourhoods. These postcode sectors therefore, create somewhat arbitrary, discrete boundaries around an individual and the analysis doesn't allow for the fact that individuals may also be influenced by processes occurring in neighbouring postcode sectors. Furthermore, the postcode sectors are not centred on each individual so this issue may be particularly salient where individuals live towards the edge of their postcode sector. These limitations mean that it is likely that an individual's neighbourhood context is measured with error and this error may be minimised with an improved neighbourhood design scheme (Owen et al., 2016).

Additionally as there is a lack of data on residential history, each individual is assigned to one neighbourhood at one point in time. It is quite likely, however, that each person has moved neighbourhood in the past, so in reality any neighbourhood effect will come from a combination of neighbourhoods, not just the neighbourhood in which they currently reside (Wodtke et al., 2011). Furthermore even if they have not changed neighbourhoods it is likely that the neighbourhoods themselves have changed. Further research into gene-environment interactions should therefore look to include a longitudinal component to account for changing social environments, particularly as the social environment has become almost universally more obesogenic over time.

A further limitation is that the sample is over 90% female and therefore even though some males are included in the study, care must be taken when generalising the findings of the study to males. If the analysis is run with the males excluded the results are qualitatively the same as in the analysis presented above, however if the analysis is run with the females excluded the sample size is not

large enough to have confidence in the results. Previous research has shown different relationships between neighbourhood deprivation and obesity in males compared with females (Stafford et al., 2010) and therefore it may also be the case that the effect of the neighbourhood environment on heritability is different for males and females.

The results also highlight the limitations of using twins to infer information on genetics. As can be seen from Figure 2 the estimates for the shared environmental variance are negative between deprivation scores of 1 and 3. While it is theoretically possible that shared environments could make twins less similar, for example perhaps twins may act specifically to differentiate themselves from their co-twins, the idea that the net effect of the shared environment that twins experience makes them more different rather than similar seems somewhat unlikely.

Although it is not possible to be entirely sure what is causing this negative shared environmental variance given the data available, one plausible solution may be that it is due to the violation of one or more of the assumptions of the method. For example, it may be that the equal environments assumption is violated, due perhaps, to people treating MZ twins more similarly than DZ twins because of their greater physical similarity. If this was the case and if it manifested itself in an effect on their BMI this would lead to MZ twin correlations being inflated compared to DZ twin correlations. The heritability estimates would therefore be biased upwards and estimates of the shared environmental variation would be biased downwards. Given the available data it is not possible to tell whether this is happening and if so, by how much. Even if this is the case, however, this study is interested in the direction and magnitude of the gene-environment interaction rather than in calculating specific heritability estimates, so the above would only be an issue if the assumption was broken more strongly in some environments then others.

Moving forward replicating this analysis with other techniques is one avenue of future research. A method known as genome wide complex trait analysis (GCTA) has recently been developed as a way of assessing a lower bound for heritability. It estimates the total additive genetic influence due to

single nucleotide polymorphisms (SNPs) across the whole genome and has capacity for dealing with genome wide interactions (Yang et al., 2011). SNPs are variation at a single position in the DNA sequence and are the most common form of genetic variation between individuals. As it uses molecular genetic data this method has the distinct advantage of not making the assumptions related to twin studies. A possible barrier to the use of this and other genome wide methods for studying gene-environment interactions is issues of personal data protection which arise from the requirement of genetic data alongside data on residential location of individuals. Furthermore studies using genome wide data require large sample sizes (Boardman et al., 2014).

In conclusion this study, despite the limitations outlined, is one of the first to include genetics into research on obesogenic environments and has shown the significance and importance of so doing. Future research is needed to build upon these initial findings, using improved measures of both genetics and the environment.

References

Ahmad, S., Rukh, G., Varga, T. V., Ali, A., Kurbasic, A., Shungin, D., Ericson, U., Koivula, R. W., Chu, A. Y., Rose, L. M., Ganna, A., Qi, Q. B., Stancakova, A., Sandholt, C. H., Elks, C. E., Curhan, G., Jensen, M. K., Tamimi, R. M., Allin, K. H., Jorgensen, T., Brage, S., Langenberg, C., Aadahl, M., Grarup, N., Linneberg, A., Pare, G., Magnusson, P. K. E., Pedersen, N. L., Boehnke, M., Hamsten, A., Mohlke, K. L., Pasquale, L. T., Pedersen, O., Scott, R. A., Ridker, P. M., Ingelsson, E., Laakso, M., Hansen, T., Qi, L., Wareham, N. J., Chasman, D. I., Hallmans, G., Hu, F. B., Renstrom, F., Orho-Melander, M., Franks, P. W., InterAct Consortiutm & DIRECT Consortium. (2013). Gene x Physical Activity Interactions in Obesity: Combined Analysis of 111,421 Individuals of European Ancestry. *Plos Genetics,* 9.

Beckwith, J., & Morris, C. A. (2008). Twin studies of political behavior: Untenable assumptions?. *Perspectives on Politics*, *6*(04), 785-791.

Boardman, J. D., Daw, J., & Freese, J. (2013). Defining the environment in gene–environment research: Lessons from social epidemiology. *American journal of public health*, *103*(S1), S64-S72.

Boardman, J. D., Domingue, B. W., Blalock, C. L., Haberstick, B. C., Harris, K. M., & McQueen, M. B. (2014). Is the gene-environment interaction paradigm relevant to genome-wide studies? The case of education and body mass index. *Demography*, *51*(1), 119-139.

Boardman, J. D., Roettger, M. E., Domingue, B. W., McQueen, M. B., Haberstick, B. C., & Harris, K. M. (2012). Gene–environment interactions related to body mass: School policies and social context as environmental moderators. *Journal of Theoretical Politics*, *24*(3), 370-388.

Cummins, S., & Macintyre, S. (2006). Food environments and obesity—neighbourhood or nation?. *International journal of epidemiology*, *35*(1), 100-104.

Cummins, S. C., McKay, L., & MacIntyre, S. (2005). McDonald's restaurants and neighborhood deprivation in Scotland and England. *American journal of preventive medicine*, *29*(4), 308-310.

Diez Roux, A. V., & Mair, C. (2010). Neighborhoods and health. *Annals of the New York Academy of Sciences*, *1186*(1), 125-145.

Egger, G., & Swinburn, B. (1997). An" ecological" approach to the obesity pandemic. *BMJ: British Medical Journal*, *315*(7106), 477.

Elks, C. E., Den Hoed, M., Zhao, J. H., Sharp, S. J., Wareham, N. J., Loos, R. J., & Ong, K. K. (2012). Variability in the heritability of body mass index: a systematic review and meta-regression. *Frontiers in endocrinology*, *3*, 29.

Flegal, K. M., Kit, B. K., Orpana, H., & Graubard, B. I. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *Jama*, *309*(1), 71-82.

Galster, G. C. (2012). The mechanism (s) of neighbourhood effects: Theory, evidence, and policy implications. In *Neighbourhood effects research: New perspectives* (pp. 23-56). Springer Netherlands.

Keller, M. C. (2014). Gene× environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biological psychiatry*, *75*(1), 18-24.

Kopelman, P. (2007). Health risks associated with overweight and obesity. *Obesity reviews*, *8*(s1), 13-17.

Lake, A., & Townshend, T. (2006). Obesogenic environments: exploring the built and food environments. *The Journal of the Royal society for the Promotion of Health*, *126*(6), 262-267.

Office for National Statistics. (2015) Harmonised Concepts and Questions for Social Data Sources. Primary principles. *4.5*

Owen, G., Harris, R., & Jones, K. (2016). Under examination: Multilevel models, geography and health research. *Progress in Human Geography*, *40*(3), 394-412.

Pampel, F. C., Krueger, P. M., & Denney, J. T. (2010). Socioeconomic disparities in health behaviors. *Annual review of sociology*, *36*, 349-370.

Pearce, J., Witten, K., Hiscock, R., & Blakely, T. (2007). Are socially disadvantaged neighbourhoods deprived of health-related community resources?. *International journal of epidemiology*, *36*(2), 348-355.

Pillinger, R., (2012). Heritability and Environmentality of WISC subscales [Online]. Available: https://www.researchgate.net/profile/Rebecca_Pillinger/publication/275333283_Heritability_and_environmentality_of_WISC_subscales/links/5537b33a0cf268fd0018c3c2.pdf [Accessed 18/12/2016].

Qi, Q., Chu, A. Y., Kang, J. H., Huang, J., Rose, L. M., Jensen, M. K., Liang, L. M., Curhan, G. C., Pasquale, L. R., Wiggs, J. L., De Vivo, I., Chan, A. T., Choi, H. K., Tamimi, R. M., Ridker, P. M., Hunter, D.J., Willet, W. C., Rimm, E. B., Chasman, D. I., Hu, F. B. & Qi, L. (2014). Fried food consumption, genetic risk, and body mass index: gene-diet interaction analysis in three US cohort studies. *Bmj*, *348*, g1610.

Rabe-Hesketh, S., Skrondal, A., & Gjessing, H. K. (2008). Biometrical modeling of twin and family data using standard mixed model software. *Biometrics*, *64*(1), 280-288.

Rosenquist, J. N., Lehrer, S. F., O'Malley, A. J., Zaslavsky, A. M., Smoller, J. W., & Christakis, N. A. (2015). Cohort of birth modifies the association between FTO genotype and BMI. *Proceedings of the National Academy of Sciences*, *112*(2), 354-359.

Shanahan, M. J., & Hofer, S. M. (2005). Social context in gene–environment interactions: Retrospect and prospect. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(Special Issue 1), 65-76.

Stafford, M., Brunner, E. J., Head, J., & Ross, N. A. (2010). Deprivation and the development of obesity: a multilevel, longitudinal study in England. *American journal of preventive medicine*, *39*(2), 130-139.

Swinburn, B., Egger, G., & Raza, F. (1999). Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Preventive medicine*, *29*(6), 563-570.

Townshend, T., & Lake, A. A. (2009). Obesogenic urban form: theory, policy and practice. *Health & place*, *15*(4), 909-916.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological science*, *14*(6), 623-628.

Tuvblad, C., Grann, M., & Lichtenstein, P. (2006). Heritability for adolescent antisocial behavior differs with socioeconomic status: gene–environment interaction. *Journal of Child Psychology and Psychiatry*, *47*(7), 734-743.

van Lenthe, F. J., Brug, J., & Mackenbach, J. P. (2005). Neighbourhood inequalities in physical inactivity: the role of neighbourhood attractiveness, proximity to local facilities and safety in the Netherlands. *Social science & medicine*, *60*(4), 763-775.

Wodtke, G. T., Harding, D. J., & Elwert, F. (2011). Neighborhood effects in temporal perspective: The impact of long-term exposure to concentrated disadvantage on high school graduation. *American Sociological Review*, *76*(5), 713-736.

World Health Organisation. (2015). *Controlling the global obesity epidemic* [Online]. Available: http://www.who.int/nutrition/topics/obesity/en/ [Accessed 03/07/2015].

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, *88*(1), 76-82.

Figure captions

Figure 1 Estimated heritability with respect to neighbourhood deprivation

Figure 2 Estimated variances attributed to each type of variation with respect to neighbourhood deprivation

Tables

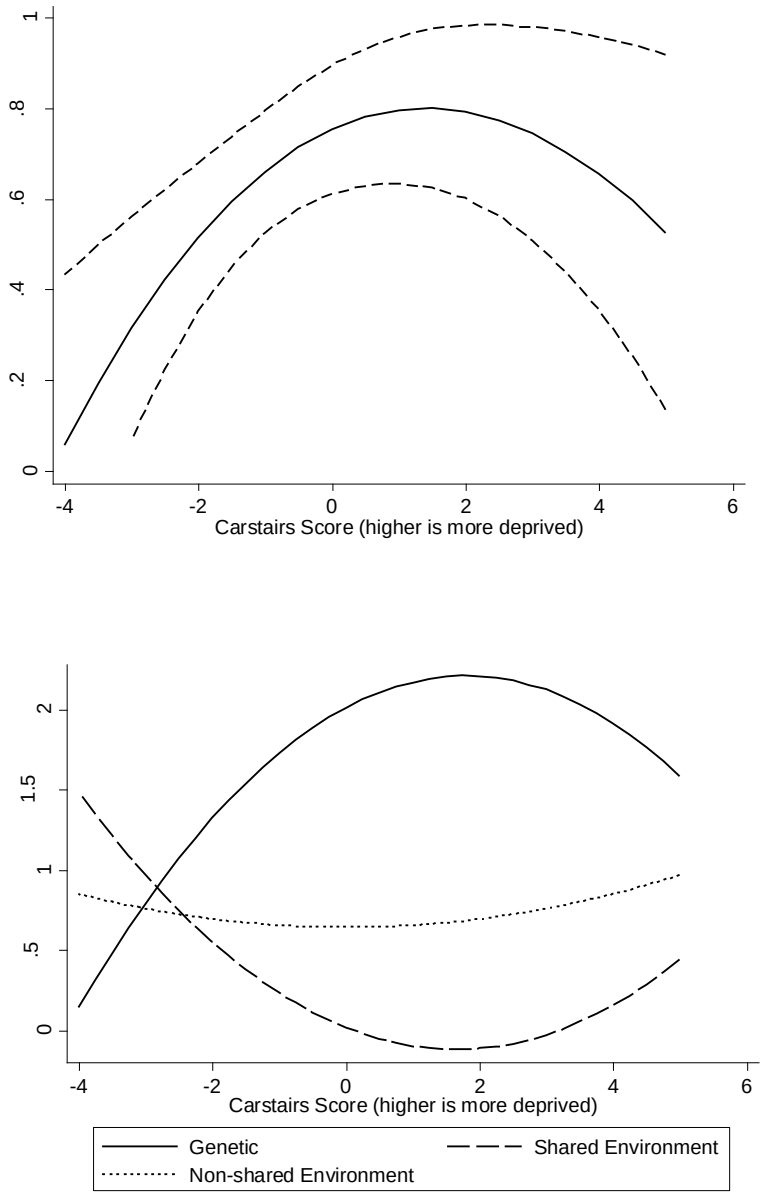
Table 1 Summary statistics and information about the variables

Variable	Summary Statistics/categories	Notes
BMI	Min 14.75 Max 52.39 Mean 25.37 Median 24.50	Calculated from measurements of height and weight. Log transformed in the models due to non-normality in the residuals and then multiplied by ten to aid interpretability of results.
Age	Min 16.11 Max 79.73 Mean 46.77 Median 48.36	A quadratic term is included alongside the linear age term because of the non-linearity in the relationship between age and BMI. Mean centred in all models to give interpretable intercept and to help estimation.
Sex	Female (n = 2,872) Male (n = 256)	
Ethnicity	White (n = 3,090) Black (n = 14) Asian (n = 8) Chinese (n = 2) Mixed (n = 8) Other (n = 6)	Where only one twin's ethnicity was recorded the twin with no ethnicity was attributed the same ethnicity as their co-twin.
Education	No qualifications (n = 435) Other qualifications (n = 134) Below GCSE level (n = 99) GCSE level and equivalents (n = 693) A levels and equivalents (n = 435) Higher education and equivalents (n = 1332)	Education data was taken from multiple questionnaires which all asked variations on the question 'what is your highest educational qualification?' The possible question response categories varied, therefore each categorisation was transferred into classifications used by the Office for National Statistics (2015). GCSEs are typically taken at age 16, while A levels are typically taken at age 18.
Deprivation	Min -5.36 Max 16.33 Mean -0.40 Median -1.01	Mean centred in all models to give interpretable intercept and to help estimation.
Year	Oldest 1992 Most Recent 2007 Mean 2000 Median 1999	Year in which measurement was taken centred round the year 2000.

Table 2 Parameter estimates for all models. Standard errors in parentheses

		Model 1	Model 2	Model 3
Random Parameters				
Genetic	σ^2_{a0}	1.853 (1.490 to 2.216)	1.860 (1.500 to 2.220)	2.001 (1.615 to 2.387)
	σ_{a01}			0.111 (0.029 to 0.193)
	σ^2_{a1}			-0.061 (-0.022 to -0.100)
Shared Environmental	σ^2_{c0}	0.149 (-0.190 to 0.488)	0.102 (-0.237 to 0.441)	0.018 (-0.343 to 0.379)
	σ_{c01}			-0.083 (0.003 to 0.163)
	σ^2_{c1}			0.050 (0.013 to 0.087)
Non-shared Environmental	σ^2_{e0}	0.691 (0.620 to 0.762)	0.695 (0.624 to 0.766)	0.641 (0.561 to 0.721)
	σ_{e01}			-
	σ^2_{e1}			0.013 (0.001 to 0.025)
Fixed Parameters				
Constant		30.378 (29.864 to 30.892)	32.066 (31.952 to 32.180)	32.073 (31.959 to 32.187)
Age		0.083 (0.065 to 0.101)	0.084 (0.066 to 0.102)	0.084 (0.066 to 0.102)
Age^2		-0.000 (-0.000 to -0.000)	-0.000 (-0.000 to -0.000)	-0.000 (-0.000 to -0.000)
Sex Reference - Female	Male	0.239 (-0.020 to 0.498)	0.272 (0.015 to 0.529)	0.269 (0.012 to 0.526)
Year		0.059 (0.041 to 0.077)	0.063 (0.045 to 0.081)	0.062 (0.044 to 0.080)
Education Reference - Higher Education and professional/ vocational equivalents	A levels equivalents		0.132 (-0.017 to 0.281)	0.121 (-0.026 to 0.268)
	GCSE level and equivalents		0.111 (-0.026 to 0.248)	0.111 (-0.024 to 0.246)
	Below GCSE level		0.152 (-0.122 to 0.426)	0.137 (-0.135 to 0.409)
	Other Qualifications		0.192 (-0.071 to 0.455)	0.174 (-0.089 to 0.437)
	No Qualification		0.340 (0.160 to 0.520)	0.319 (0.141 to 0.497)
Ethnicity Reference - White	Mixed		-0.501 (-1.608 to 0.606)	-0.572 (-1.697 to 0.553)
	Black		0.293 (-0.748 to 1.334)	0.273 (-0.803 to 1.349)
	Asian		-0.247 (-1.605 to 1.111)	-0.211 (-1.593 to 1.171)
	Chinese		-0.925 (-3.910 to 2.060)	-0.901 (-3.900 to 2.098)
	Other		0.408 (-0.958 to 1.774)	0.285 (-1.032 to 1.602)
Deprivation Score			0.021 (0.001 to 0.041)	0.026 (0.006 to 0.046)
-2*loglikelihood (IGLS Deviance)		11282.7	11261.3	11234.5

Figures


0
.2
.4
.6
.8
1
-4
-2
0
2
4
6
Carstairs Score (higher is more deprived)
0
.5
1
1.5
2
-4
-2
0
2
4
6
Carstairs Score (higher is more deprived)
Genetic
Shared Environment
Non-shared Environment